



# A Revised Approach to IT Modernization



*Centralized Revenue Opportunity System (CROS)*

---

State of California  
Board of Equalization  
450 N Street  
PO Box 942879, MIC: 01

*Eric Steen, CROS Project Director  
April 2013*

## Contents

Introduction .....	1
Purpose .....	2
Audience .....	3
PI-I Core Practices .....	3
Program Area Readiness .....	3
Data Readiness.....	4
Interfaces .....	6
System Inventory .....	7
Final Remarks.....	8
References .....	9

# A Revised Approach to IT Modernization

---

## Introduction

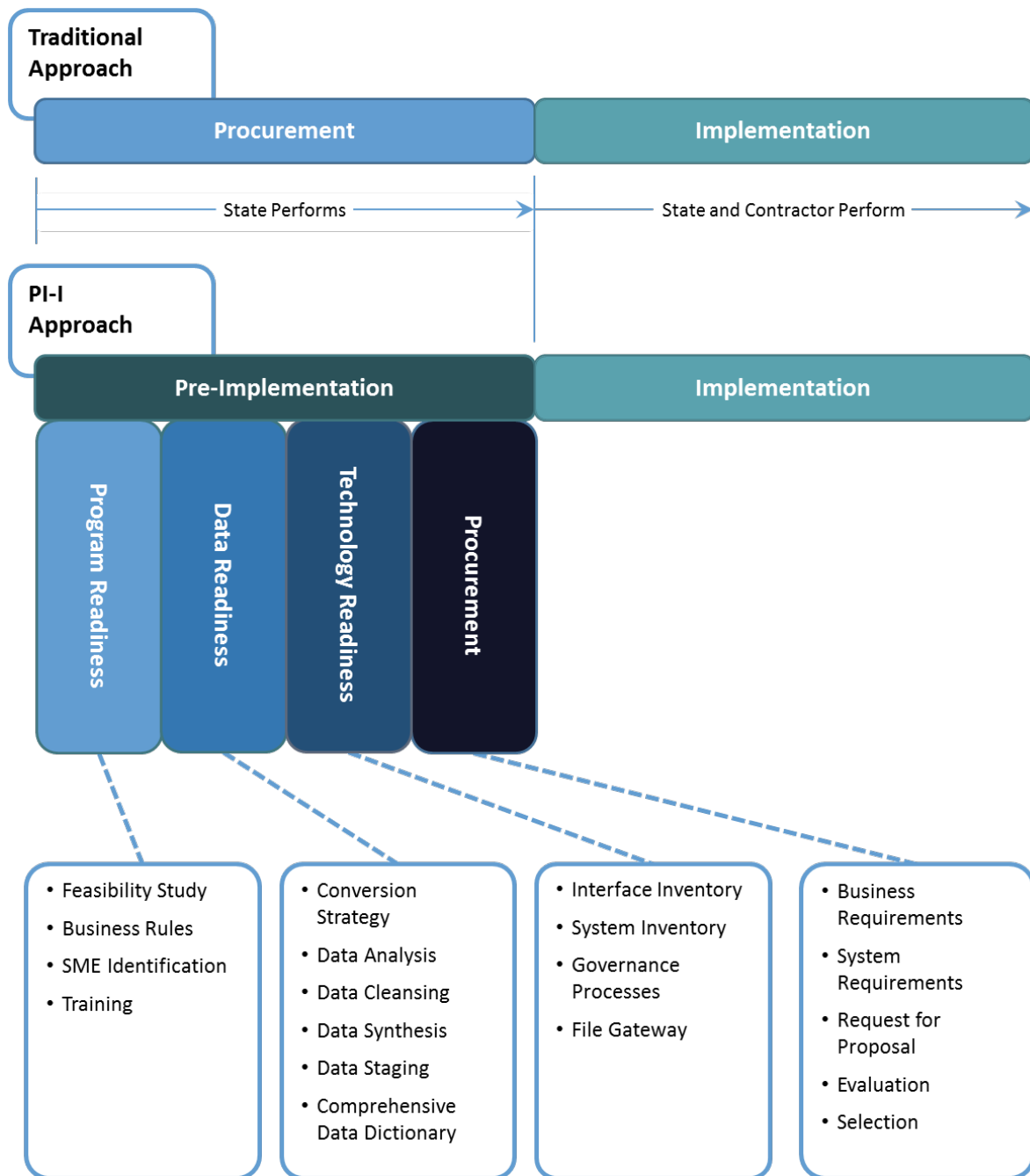
The California State Board of Equalization (BOE) is undertaking an Information Technology (IT) modernization effort entitled the Centralized Revenue Opportunity System (CROS). The project's primary objectives are to improve the taxpayer experience and internal operations. It will also improve revenue collection by providing a vehicle to detect and resolve underreporting.

The backdrop for BOE stakeholders – legislators, control agencies, and the taxpayer – is California's track record when it comes to such undertakings. In many respects California is the cradle of IT civilization. Silicon Valley is the home of Apple, Google, Facebook, and numerous startups. Ironically, the State falters when it comes to delivering IT projects in the public sector.

As a result of these challenges, California is examining its IT procurement approach. The current procurement process, designed to ensure a level playing field among bidders and protect the State's interests, is lengthy. Consequently, procurements take many years. In some cases proposed technologies are outdated by the time a vendor signs a contract with the State. Regardless of whether California streamlines the procurement cycle, projects will continue to face implementation risk. Therefore, the BOE has adopted an approach that invests in critical activities before the IT project's implementation and along with procurement. This paper highlights BOE's activities, which any department can undertake to reduce project risk and that do not rely on changes in law, regulation, or outside governance.

The procurement phase itself does little to prepare the organization for implementation. Instead, projects should be viewed as *pre-implementation* and *implementation* (referred to as PI-I throughout), wherein pre-implementation includes the procurement and other activities critical to solution delivery. The PI-I approach requires more State resources and upfront costs. However, PI-I reduces overall costs by mitigating significant risk and shifting tasks normally conducted by the contractor to the State, where the work ultimately belongs.

Pre-implementation uncovers issues that commonly derail projects – poor data quality, lack of subject matter expertise, undocumented interfaces – early rather than later. It also gives State employees greater involvement in, and ownership of, the project. When proven successful, the BOE hopes that this approach will be used as a model for similar efforts throughout the State. A graphic depicting PI-I is featured below.



## Purpose

This paper is based upon concrete actions taken by the BOE that can be applied to other projects. This document explains common risks, explains BOE's approach to handling them (almost like mini-case studies), and then provides recommendations that others may apply. The BOE hopes that PI-I proves successful and is adopted as a model for future projects and that the recommendations within are incorporated into future feasibility studies.

## Audience

This paper is intended for those who want to understand how the BOE is working to ensure that the CROS Project succeeds. Public-sector executives and managers contemplating a large-scale IT modernization project will find this information beneficial. Other interested parties may include the Legislature and control agencies, which are the budgetary, spending, and oversight authorities. The document is written for a general audience, but provides enough technical detail to be actionable.

## PI-I Core Practices

The core practices addressed in the following sections are commonly performed by the contractor on large projects. These practices reveal problems that, if not resolved early, will delay implementation. The premise of PI-I is that State staff handle these activities in concert with procurement, minimizing risks and lowering costs. Furthermore, by taking on these challenges, State resources reinforce project skills so that they are prepared and self-confident when it's time for implementation.

## Program Area Readiness

IT modernization's objective is to improve the organization's ability to achieve its business goals more efficiently and at lower cost. While much of the work has to do with technology, implementation cannot happen without Subject Matter Experts (SMEs) from within the organization's program areas. This expertise is vital to articulating functional requirements, refining business processes, improving policies and practices, and validating that the final solution meets business objectives.

Frequently in large organizations, business processes and rules that govern requirements are poorly documented. Documentation is difficult to understand, contradictory, or nonexistent. This information is critical to crafting a good solution. The BOE took great pains to involve program area experts to document the business needs within the scope of CROS. This work is also integral to cleansing data in preparation for conversion.

The State should identify and involve SMEs early in the project lifecycle, long before the contractor arrives. Engaging SMEs early ensures that the necessary documentation is complete and increases State staff's involvement in and ownership of the project. This approach also lowers costs. Since State staff have largely completed this work, the contractor will have less to do in this area. Moreover, these SMEs become the liaisons to the contractor and assist them with requirements and design.

## Recommendations

- Communicate to the organization that the project is an enterprise goal, and that comprehensive participation is integral to the State's success.
- Identify SMEs from every relevant program area and involve them to craft the Request for Proposal (RFP) and serve as liaisons to the contractor.
- As part of the RFP's functional requirements, document business processes using process-flow models.
- Document the detailed business rules that govern functionality and data.

## Data Readiness

The contractor's responsibility is to realize the vision described in the RFP. This vision commonly includes a new application that handles core business functions, which may be custom built, off-the-shelf, or some combination of the two. The application may be accompanied by support capabilities such as a data warehouse, integration with other systems, network enhancements, reengineered business processes, and so on. However, one thing is certain: the new system will require data from the legacy systems, which means data conversion.

Projects often falter because of the inability to convert data from legacy assets. There is no benefit to standing up "shiny new systems" if you cannot effectively get your data into them.<sup>i</sup> Typically, data conversion is relegated to the contractor, which means that the effort starts following procurement. There are two problems with this approach. First, the effort commences too late as conversion problems are typically complicated and time consuming to address. Second, data readiness (a prerequisite to conversion) should be the responsibility of the State, as it knows its data better than anyone. Readiness entails assessing data quality, cleansing anomalies, synthesizing duplicates, staging, and documenting data definitions. Once the data is cleansed and staged, the contractor can map data from the staging area to the target environment, craft conversion scripts, and then extract, transform, and load the data into the target environment.

BOE's environment is composed of two primary systems and a host of ancillary systems. One system uses an ADABAS database; the other uses Sybase SQLServer. Together they house more than 500 gigabytes of data. The first step in the readiness process was to identify the functional areas within which data could be categorized, e.g., client registration, revenue, payments, delinquencies. Following the identification of functional areas, the BOE began to assess the data quality for each functional area. Because the ADABAS system is the system of record it was addressed first. The immediate challenge BOE faced was loading production data into an environment where it could be assessed, as the production environment was not easily accessible. Staff crafted scripts that transformed hierarchical metadata into relational structures implemented in a Microsoft SQLServer staging repository. Data was then extracted from ADABAS, transformed, and then applied to the relational environment.

The rationale for migrating data to a relational staging repository is threefold:

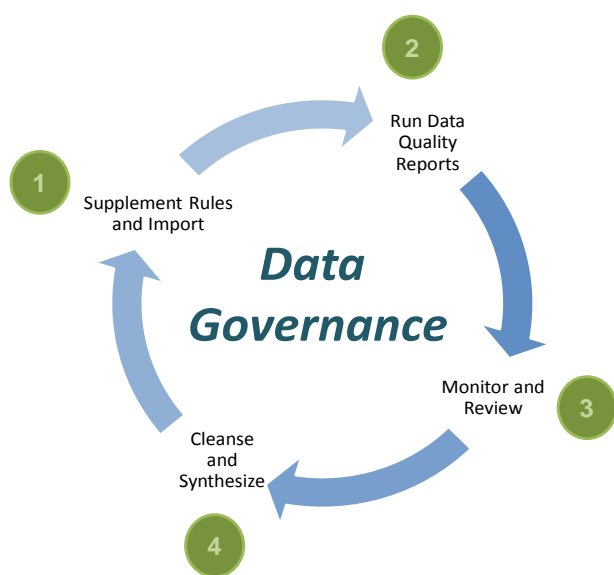
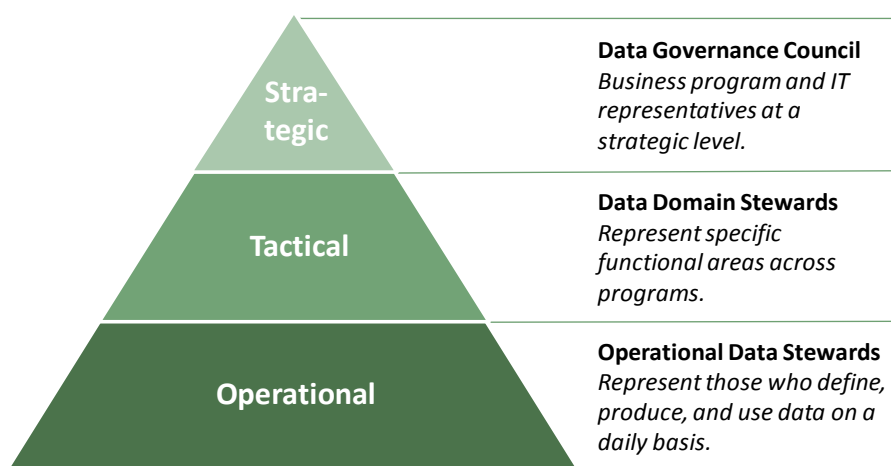
1. All data sources had to be synthesized in a single repository, and a relational database made it easier to facilitate this effort.
2. State data center costs are based on processor utilization, which substantially drives up the cost to conduct quality assessment. Moving data into BOE's internal environment significantly reduced costs.
3. The data quality tools relied on a relational database.

To assist BOE SMEs with data quality assessment, the BOE hired a small, experienced consulting firm that utilized its custom-developed data quality tools. BOE SMEs explained policies and practices governing data, which were input into these tools. Application programming rules embedded in legacy systems were also entered into these tools. Once quality rules were in place, the tools evaluated the

legacy data residing in the staging repository. The tools generate reports that highlight data quality violations and set in motion efforts to remedy problems.

A critical aspect of this effort is data governance, which prescribes the authority and control over data fixes – deciding which data is authoritative, how fixes might apply, standards to which data must adhere, and who applies fixes and when they are applied. Since State program area experts are the most qualified to make such decisions, this is not a process that was outsourced to consultants or to technical staff.

At the BOE, data stewards settle issues pertaining to data quality, security, access, etc. Decisions pertaining to standards and synthesizing duplicated data across programs are facilitated through the Data Governance Council. The Data Governance Council includes managers and executives representing all business areas within the scope of CROS. The following graphic depicts this governance model.



The process of reviewing and fixing data is iterative since production data grows continually – new businesses emerge, returns are filed, and so on. (The BOE implemented real-time replication to dynamically capture updates in the staging repository.) Periodic data quality reports also reveal the pace of cleansing, at times indicating the need to adjust resources. This lather-rinse-repeat cycle eventually results in a clean data repository ready for conversion. (The repository also supports test data creation and business intelligence activities.) While the effort to extract, transform, and load legacy data to the new environment remains, this exercise becomes easier by using the staging repository and a thorough data dictionary.

## Recommendations

- Identify all data sources targeted for conversion.
- Only hire outside staff with proven experience in data conversion when experienced internal resources are unavailable.
- Employ data quality tools to assess legacy data. (Many tools are commercially available and vary in price. Tier-1 vendors have robust offerings, but open source alternatives are available, too. The following papers may prove helpful: *Measuring the Business Value of Data Quality*<sup>ii</sup> and *Magic Quadrant for Data Quality Tools*<sup>iii</sup>.)
- Establish an internal data governance team composed of SMEs, and empower them to make decisions regarding data cleanup.
- Provide a data repository containing clean, synthesized data.
- Craft a well written and exhaustive data dictionary that thoroughly describes legacy data.

## Interfaces

Every organization exchanges data, internally and externally. The number of exchanges grows over time, and rarely is there a central and authoritative source for this information. Often there is no single group that handles how interagency agreements (IAA) or memoranda of understanding (MOU) are brokered. Data is exchanged using various media – email, File Transfer Protocol (FTP), and even hand-delivered on compact discs. Seldom is there a single group that maintains the tools supporting electronic data transfer. And because formats are varied and inconsistent, numerous parties may be responsible to massage inbound or outbound data to make it usable. This lack of uniformity makes it difficult to assess the scope of the interface effort. Undoubtedly, the target solution must maintain existing exchanges, so if the scope is poorly defined, the organization is likely to face increased costs, schedule slippage, and possibly challenges to implementation.

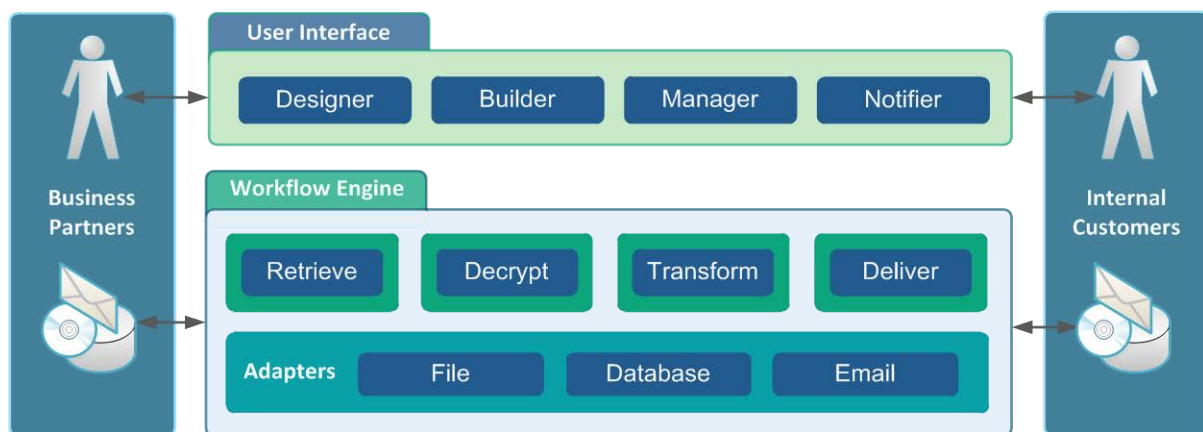
Like most organizations, the BOE exchanges a high volume of information with partners. While internal file sharing will be reduced by legacy replacement, exchanges with external parties will not. To date, the BOE has identified roughly 500 interfaces, a third of which are central to CROS. The BOE has begun documenting each of these interfaces to include the purpose, frequency, edge systems involved, medium (e.g., email, mail, FTP), direction (inbound or outbound), contact information, file layout, sample data, security constraints, and associated IAAs or MOUs. The BOE has also documented data transformation rules (i.e., data format requirements) when appropriate.

In addition to cataloging all interfaces, the BOE is taking steps to create a centralized service that orchestrates enterprise data exchange. The BOE embarked on this approach based on management's experience on similar projects and control agency stipulations that BOE improve data sharing with the other revenue agencies. Moreover, this decision is supported by lessons learned at other State agencies. This undertaking ensures that exchange occurs in a consistent and secure way and makes it easier for external partners to automate file submission and acquisition.

The BOE's file gateway supports all file-based transactions and has the ability to transform data dynamically (e.g., fixed-length, delimited, XML). This effort requires working closely with partners to



modify how files are shared; in some cases, agreements must be forged regarding security standards, protocols, and file formats. The following high-level diagram depicts the data exchange gateway.



The gateway provides a web-based user interface that enables administrators to configure data exchange. The administrator can select from a variety of adapters to support the desired communication method such as email or FTP. Administrators can then chain together activities in workflows to acquire a file, archive a copy, decrypt it if needed, transform the data, and put data in a file directory or apply data directly to a system. The gateway also provides features to monitor the exchange, identify where files are in the workflow, and inspect and handle errors.

### Recommendations

- Catalog all interfaces to include details such as descriptions, file layout, system endpoints.
- Collect all supporting documentation such as IAAs, MOUs, and sample data files.
- Build a cohesive, one-stop-shop responsible for brokering interface agreements and supporting automated data exchange.
- Establish standards (e.g., security requirements, protocols, and templates) for how interfaces should be documented.
- Consider implementing a Business to Business (B2B) gateway to automate exchange with partners.

### System Inventory

The genesis of every modernization effort is deficient legacy systems; were existing infrastructure sufficient, there would be no reason to invest heavily in new hardware, software, and implementation services. And where deficiencies exist, workarounds also exist. This means that ancillary solutions – manual processes or one-off technology implementations – surround kernel systems. One-off technology solutions may be mission critical to program areas but unsupported by the technology department. Left unaccounted for, these systems are a risk to project scope; the new solution may lack mission-critical capabilities or fail to interface with remaining legacy systems.

At BOE these ad hoc systems have been implemented using Access, Excel, and Delphi, and other technologies. The CROS team spent months tracking down and documenting these systems to

determine if their data and functionality should be incorporated into the RFP requirements for the target solution. For those not slated for replacement, the BOE had to determine whether the CROS solution would interface with them.

### **Recommendations**

- Document all systems in scope for replacement. Include a functional description, platform description, and a description of data residing within the system. Ensure that manual systems are not overlooked during this exercise.
- Decide the future of each system:
  - Retain the system and determine whether it must interface with the target solution.
  - Replace the system and ensure that its functionality is incorporated into the solution.
  - For systems slated for replacement, determine whether the data must be migrated to the target solution.
- Communicate the inventory and disposition to the entire organization and request signoff that the inventory is complete. Signoff reduces the likelihood that a system is left off the list only to later emerge and negatively impact scope.

### **Final Remarks**

Because core practices are critical to successful implementation, the PI-I approach requires that an organization dedicate its best resources to a project and leverage in-house expertise, something that often isn't done because of competing priorities. In BOE's case, senior leadership realigned priorities to ensure that CROS was staffed properly with in-house program and technology experts who have the skills to select and apply the right tools to meet business objectives.

An agency is ultimately responsible for a project's success. The public rightly expects its civil servants to deliver successful outcomes. If executed properly, PI-I lowers costs, mitigates risk, accelerates progress, and invests in State staff. In-sourcing up front tasks like data readiness and external interfaces facilitates organizational buy-in, initiates the cultural change needed to embrace a solution, and seeks to avoid long-term dependency on contractors. The end result is a modernized IT solution that meets the agency's business objectives and fosters the organizational change needed to maximize the solution's capabilities.

## References

---

<sup>i</sup> John Morris, *Practical Data Migration* (Swindon, UK: The British Computer Society, 2009), xix

<sup>ii</sup> Ted Freedman, Michael Smith, “Measuring the Business Value of Data Quality,” *Gartner*, Oct 10, 2011

<sup>iii</sup> Ted Friedman, Andreas Bitterer, “Magic Quadrant for Data Quality Tools,” *Gartner*, Jul 28, 2011